Software Reuse Enablement System Trade Study

Earth Science Data Systems

Software Reuse Working Group

Contributing Working Group Members

- Victor Delnore *
- Robert R. Downs
- Ryan Gerard
- Steve Olding
- Shahin Samadi
- Robert Wolfe *
- Nancy Casey
- Stefan Falke
- David Giles
- Tommy Jasmin
- Ross Swick
- Bill Teng

* Co-chair

Trade Study Document Evolution

- Document created after performing initial research
- Support group provided feedback and revisions
- First draft sent to Reuse Working Group for feedback and comments
- Second draft sent to:
  - GCMD, ECHO, and GSFC Open Source representatives for feedback on their sections
  - Other Working Group chairs for their comments
- Final draft created
- Preparing for presentation/submission to HQ

General Background

- Software Reuse Working Group fosters software reuse across Earth science systems.
- Goal is to encourage software developers to make use of existing software assets, including open source.
- A survey on the reuse practices of the Earth science community was conducted and results show that:
  - Developers need to be able to easily locate and evaluate available reusable artifacts
  - A catalog or repository for reusable artifacts is the best means of increasing software reuse within the community
- NASA tasked the Working Group to look at the roles of the GCMD, Open Source Agreement site, and other sites in serving the community and meeting reuse needs.

Requirements Background

- Primary users of a Reuse Enablement System (RES) are NASA-funded software developers within the Earth science community who create software products.
- Working Group collaborated for several months in 2004 to identify the important functional requirements needed for such a system.
- Requirements fall into a number of areas including:
  - General
  - Asset Usage
  - Asset Submission
  - Content Management
  - System Administration

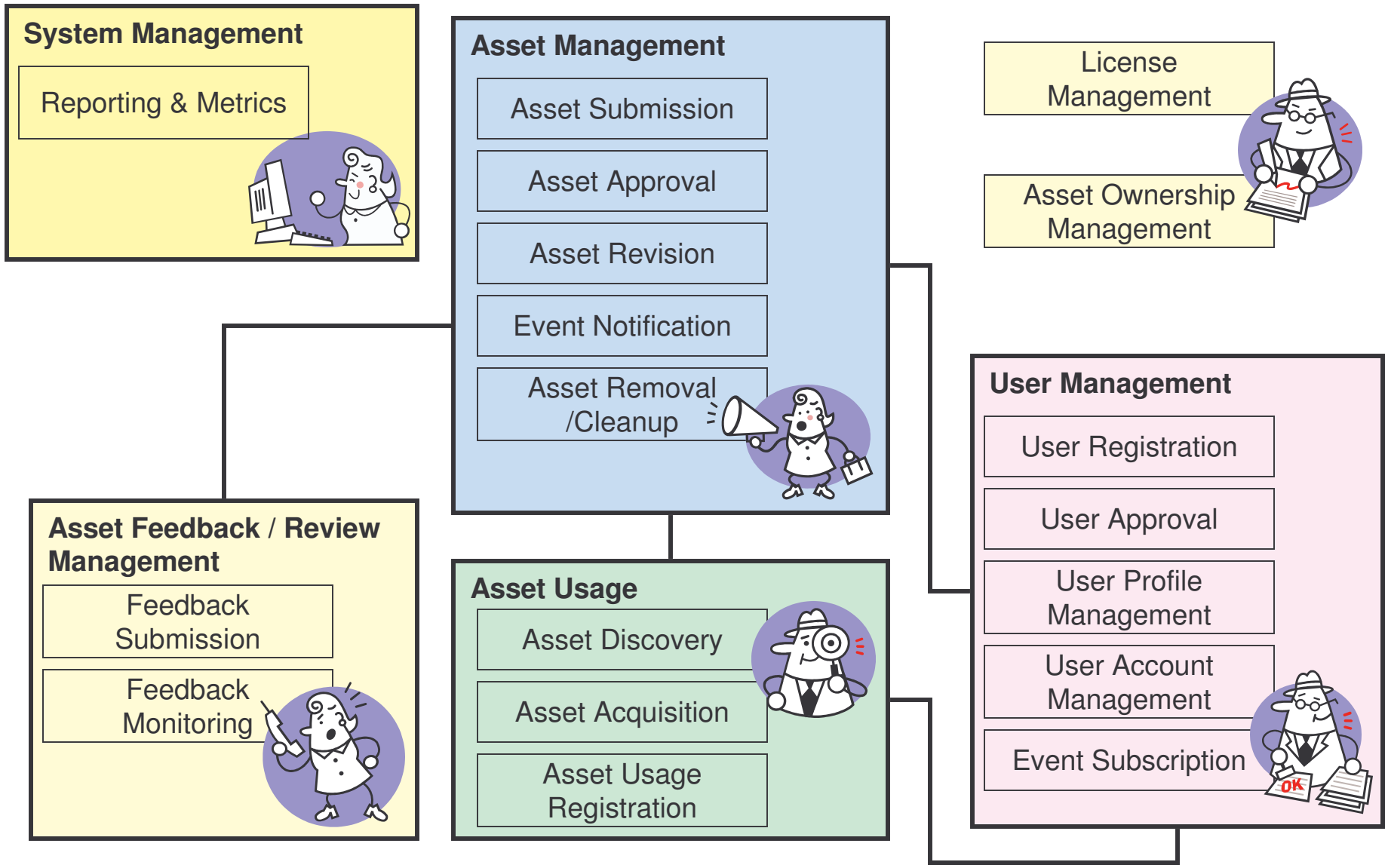
System Requirements

- Specific functional requirements identified from use cases for the system include:
  - Register User
  - Contribute/Update Assets
  - System Feedback
  - Automatic Notifications
  - Discovering Assets
  - Register Asset Usage
  - Asset Review
  - Monitoring Feedback
  - Workflow Management
  - Capture Asset Needs
  - Catalog or Repository

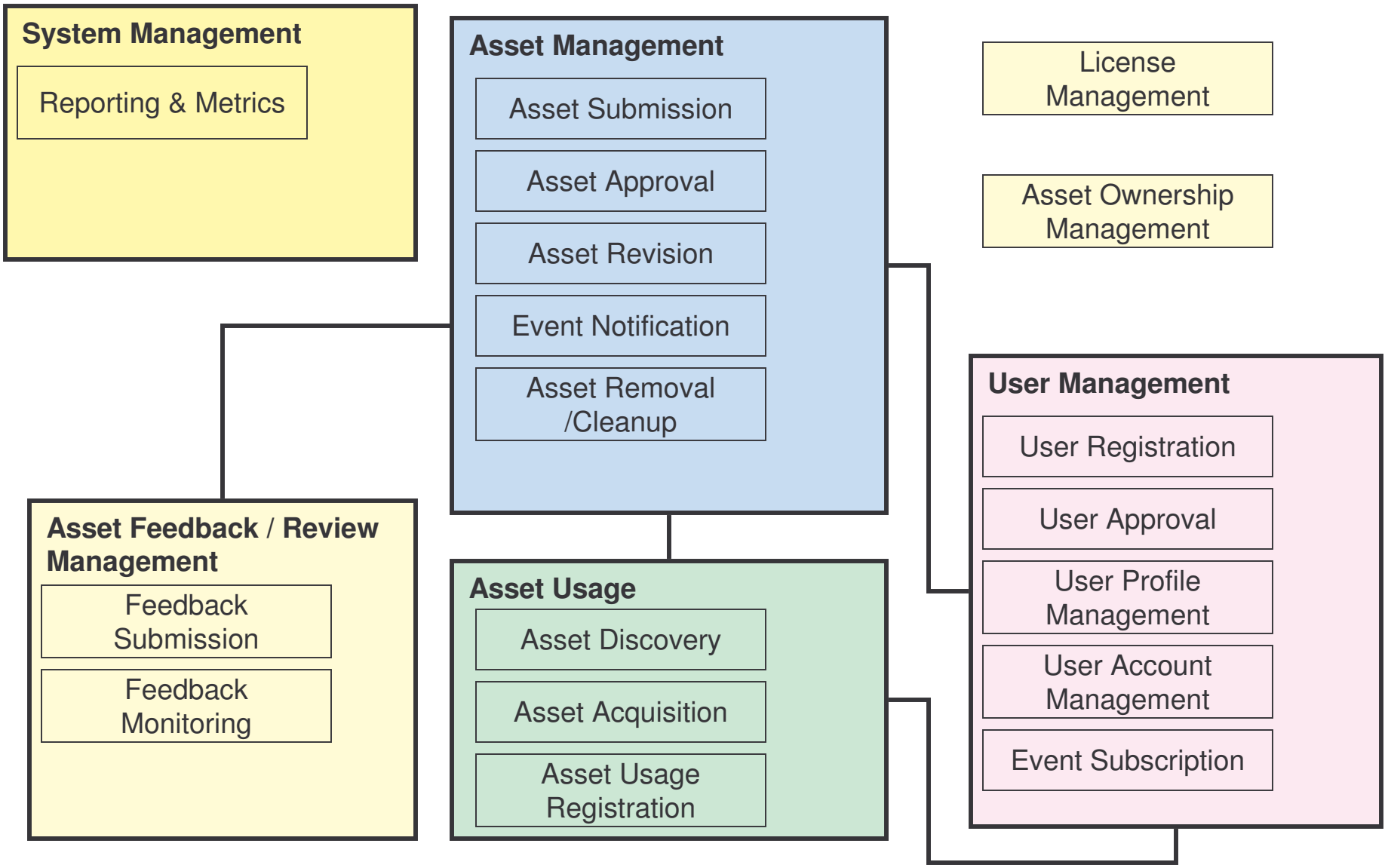
System Requirements

- Additional functional requirements include:
  - Minimal Operation Support
  - Performance
  - Security
  - Technology
- Important non-functional requirements include:
  - Domain (Earth science focus)
  - Type of assets provided (small sized components)

Relationship of Requirements



Relationship of Requirements



NASA Systems Reviewed

- NASA sites reviewed:
  - Global Change Master Directory (GCMD)
  - Goddard Space Flight Center (GSFC) Open Source Software page
  - Ames Research Center Open Source Software page
  - HDF-EOS Tools and Information Center
  - Computational Technologies (CT) Project
  - Earth Observing System Clearinghouse (ECHO)
  - Planetary Data Systems (PDS) Software Download

Non-NASA Systems Reviewed

- Non-NASA sites reviewed:
  - Open Channel Foundation (hosts NASA's COSMIC Collection)
  - SourceForge
  - Freshmeat
  - Scientific Applications on Linux
  - National Technology Transfer Center
  - National HPCC Software Exchange
  - Netlib
  - Savannah
  - Space Telescope Science Institute (STScI) Software and Hardware Products
  - Astronomical Software and Documentation Service

Other Systems Inspected

- NASA sites:
  - Direct Readout Laboratory
  - Glenn Research Center Software Repository
- Non-NASA sites:
  - ArcScripts
  - Wikipedia
  - Usenet newsgroups
  - Ruby Application Archive
  - SciRuby
  - Comprehensive Perl Archive Network
  - FreeGIS

In general, these sites were too narrowly focused to warrant a detailed review.

GCMD Site – http://gcmd.gsfc.nasa.gov/

GCMD Review Description

- Domain is Earth science
- Type of assets provided are metadata about data sets (~16000) and, to a lesser extent, data services (~1240)
- Catalog of metadata
- Operational support is from a staff of ten members which includes four developers
- System technology includes RSYNC, Zope, CVS, Linux, Java, JavaServer Pages, XML, Apache, Oracle/PostgreSQL, Struts, Lucene, XSLT, Dreamweaver
- Site registration not available
- Assets can be added and updated
- System feedback available through web form
- Automatic notifications are available and based on keywords
- Asset usage cannot be registered
- Asset reviews are not available
- System feedback is monitored
- Site does not have secure login/registration

GCMD Review Table

Requirement/Feature	*Available at GCMD?*
Domain	Earth science
Type of Assets	Data sets, data services
Register User	☆☆☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	★★★
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Catalog
Operation Support	Ten member staff Four developers
Technology	RSYNC, Zope, CVS, Linux, Java, JavaServer Pages, XML, Apache, Oracle/PostgreSQL, Struts, Lucene, XSLT, Dreamweaver

GCMD Site – http://gcmd.gsfc.nasa.gov/

- Domain is Earth science
- Type of assets provided are metadata about data sets (~16000) and, to a lesser extent, data services (~1240)
- Catalog of metadata
- Operational support is from a staff of ten members which includes four developers
- System technology includes RSYNC, Zope, CVS, Linux, Java, JavaServer Pages, XML, Apache, Oracle/PostgreSQL, Struts, Lucene, XSLT, Dreamweaver

GCMD Review

Requirement/Feature	*Available at GCMD?*
Domain	Earth science
Type of Assets	Data sets, data services
Register User	☆☆☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	★★★
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Catalog
Operation Support	Ten member staff Four developers
Technology	RSYNC, Zope, CVS, Linux, Java, JavaServer Pages, XML, Apache, Oracle/PostgreSQL, Struts, Lucene, XSLT, Dreamweaver

- Site registration not available
- Assets can be added and updated
- System feedback available through web form
- Automatic notifications are available and based on keywords
- Asset usage cannot be registered
- Asset reviews are not available
- System feedback is monitored
- Site does not have secure login/registration

GCMD Data Set Page

GODDARD SPACE FLIGHT CENTER | + Visit NASA.gov

Global Change Master Directory
a directory to Earth science data and services

About Us | FAQ | Contact Us | Site Map

Home | Data Sets | Data Services | Collaborations | Add to GCMD | What's New | Participate | Calendar | Links

- Agriculture
- Atmosphere
- Biosphere
- Climate Indicators
- Cryosphere
- Human Dimensions
- Hydrosphere
- Land Surface
- Oceans
- Paleoclimate
- Solid Earth
- Spectral/Engineering
- Sun-Earth Interactions

- Data Centers
- Locations
- Instruments
- Platforms/Sources
- Projects

- Geospatial One Stop Projects

- Free text Search

- Help Center
- Questions?

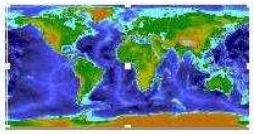
Parameters>BIOSPHERE>ANIMAL TAXONOMY>MAMMALS

Brief Record | Distribution | Attributes | Coverage | Personnel | Full Record | Update Record

Mammal Species of the World

Geographic Coverage

Southernmost Latitude: -90.0
Westernmost Longitude: -180.0
Northernmost Latitude: 90.0
Easternmost Longitude: 180.0



Summary

2005-08-12

The Mammal Species of the World (MSW) is a database of mammalian taxonomy. The names are organized in a hierarchy that includes Order, Family, Subfamily, Genus and Species. Records include the following fields: Scientific name, Author's name, and Year described, Original publication citation, Original name, Common name, Type species, Type locality, Distribution, Comments, Status. MSW contains the names of 4,629 currently recognized species of mammals. The data is from the book by Wilson, D. E., and D. M. Reeder (eds), 1993.

The list was compiled by an international team of contributors. It can be used as an on-line reference for identifying or verifying recognized scientific names and for taxonomic research; or adapted as an authority file for collections management activities of mammal collections.

Questions?

Information was obtained from http://bubl.ac.uk/link/t/taxonomy.htm and from http://www.nmnh.si.edu/msw/.

Data Center

Data Center Name: SI/NMNH >National Museum of Natural History, Smithsonian Institution
Data Center URL: http://www.nmnh.si.edu
Dataset ID: MSW

Personnel

Name: DON E. WILSON
Email: wilson.don at nmnh.si.edu
Contact Address:
National Museum of Natural History
Office of Biodiversity Programs
Smithsonian Institution
City: Washington
Province or State: DC
Country: USA

Data Set Citation

Dataset Creator: Smithsonian Institution
Dataset Title: Mammal Species of the World
Dataset Release Date: 1993
Dataset Release Place: Washington, D.C.
Dataset Publisher: Smithsonian Institution
Online Resource: http://www.nmnh.si.edu/msw/

Related URL

Content Type: BDP METADATA
URL: http://metadata.nbii.gov
Description:
Metadata in Biological Data Profile format.

FIRSTGOV Your First Click to the U.S. Government

+ Equal Employment Opportunity Data Posted Pursuant to the No Fear Act
+ Freedom of Information Act
+ FY 2004 Agency Performance and Accountability Report
+ NASA Privacy Statement, Disclaimer, and Accessibility Certification

Editor:Gene Major
NASA Official: Lola Olsen
Last Updated: September 1, 2005

Ames Open Source – http://opensource.arc.nasa.gov/

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

AMES RESEARCH CENTER

Ames Home > Open Source Software

- NOSA Software Agreement
- Other NASA Software

NASA OPEN SOURCE SOFTWARE

NASA conducts research and development in software and software technology as an essential response to the needs of NASA missions. Under the NASA Software Release policy, NASA has several options for the release of NASA developed software technologies. These options now include Open Source software release. This option is under the NASA Open Source Agreement "NOSA".

OSI certified

The motivations for NASA to distribute software codes Open Source are:

- to increase NASA software quality via community peer review
- to accelerate software development via community contributions
- to maximize the awareness and impact of NASA research
- to increase dissemination of NASA software in support of NASA's education mission

PROJECTS

Livingstone2/Skunkworks
Livingstone2 is a reusable artificial intelligence (AI) software system designed to assist spacecraft, life support systems, chemical plants or other complex systems in operating robustly with minimal human supervision, even in the face of hardware failures or unexpected events.

IND: Creation and Manipulation of Decision Trees from Data
A common approach to supervised classification and prediction in artificial intelligence and statistical pattern recognition is the use of decision trees. A tree is "grown" from data using a recursive partitioning algorithm to create a tree which (hopefully) has good prediction of classes on new data. Standard algorithms are 1) that of Breiman, Friedman, Olshen, and Stone; and 2) Id3 and its successor C4 (by Quinlan). As well as reimplementing parts of these algorithms and offering experimental control suites, IND also introduces Bayesian and MML methods and more sophisticated search in growing trees. These produce more accurate class probability estimates that are important in applications like diagnosis.

CODE
CODE is a software framework for control and observation in distributed environments. This framework enables the observation of distributed resources, services, and applications. Observations are made by modular components called sensors, the information observed is encapsulated as events, and these events are transmitted from where they are produced to whoever wants to consume them using an event management framework. Further, the CODE framework allows people or agents to control a distributed system by allowing them to take actions on remote systems using modular

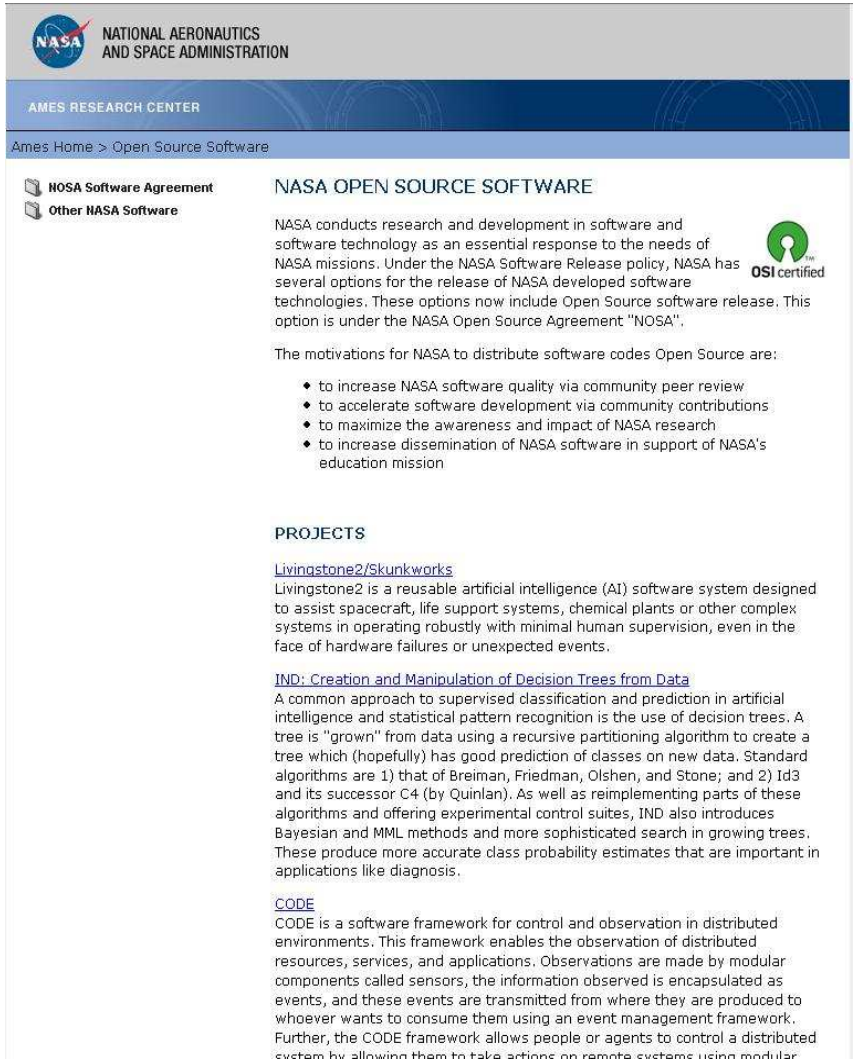
Ames Open Source Review Description

- Domain is general science
- Type of assets provided are open source packages produced by NASA (11)
- Acts as both a repository and a catalog
- Operational support is presumed to be small
- System technology is JavaServer web pages
- Site registration not available
- Assets contribution/updating not available to general public
- System feedback available through web form
- Automatic notifications are available and based on selected products
- Asset usage can be registered and is needed for notifications
- Asset reviews are not available
- System feedback is monitored
- Site does not have secure login/registration

Ames Open Source Review Table

Requirement/Feature	*Available at Ames OSS?*
Domain	General science
Type of Assets	Open source packages
Register User	☆☆☆
Contribute/Update Assets	★☆☆
Provide System Feedback	★★☆
Automatic Notifications	★☆☆
Discovering Assets	List
Register Asset Usage	★★★
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Both
Operation Support	Uncertain, presumed small
Technology	JavaServer Pages

Ames Open Source – http://opensource.arc.nasa.gov/



- Domain is general science
- Type of assets provided are open source packages produced by NASA (11)
- Acts as both a repository and a catalog
- Operational support is presumed to be small
- System technology is JavaServer web pages

Ames Open Source Review

Requirement/Feature	*Available at Ames OSS?*
Domain	General science
Type of Assets	Open source packages
Register User	☆☆☆
Contribute/Update Assets	★☆☆
Provide System Feedback	★★☆
Automatic Notifications	★☆☆
Discovering Assets	List
Register Asset Usage	★★★
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Both
Operation Support	Uncertain, presumed small
Technology	JavaServer Pages

- Site registration not available
- Assets contribution/updating not available to general public
- System feedback available through web form
- Automatic notifications are available and based on selected products
- Asset usage can be registered and is needed for notifications
- Asset reviews are not available
- System feedback is monitored
- Site does not have secure login/registration

Ames Open Source Project Page

AMES RESEARCH CENTER

Ames Home > Open Source Software

NOSA Software Agreement
Other NASA Software

WORLD WIND

[Project Home Page] [Software]

World Wind allows any user to zoom from satellite altitude into any place on Earth, leveraging high resolution LandSat imagery and SRTM elevation data to experience Earth in visually rich 3D, just as if they were really there.

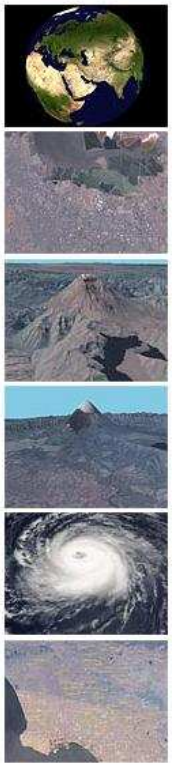
Particular focus was put into the ease of usability so people of all ages can enjoy World Wind. All one needs to control World Wind is a two button mouse. Additional guides and features can be accessed though a simplified menu. Navigation is automated with single clicks of a mouse as well as the ability to type in any location and automatically zoom into it.

The World Wind install package is all you need to get started. It contains all the other requirements such as the .NET runtime and managed DirectX library. Just download the ZIP file, extract and run the setup program.

World Wind can display a combination of data from a variety of sources...
Blue Marble – A full true color Earth as seen on NASA's Earth Observatory
LandSat 7 – An extremely detailed mosaic of imagery that's detailed enough to see freeways, stadiums, anywhere on the Earth.
SRTM – Elevation data gives rise to mountains, volcanoes, hills, and valleys.
Animated Earth – A collection of Earth science data set in motion. See how hurricanes move and fires spread.
GLOBE – See temperature, rainfall, and more across the entire globe.
Country & State borders – See outlines directly on the Earth as they trace rivers, mountain ridges, or latitude & longitude lines.

For a thorough list of features, user manual, key chart, screenshots and more, please visit http://learn.arc.nasa.gov/worldwind/

SCREENSHOTS



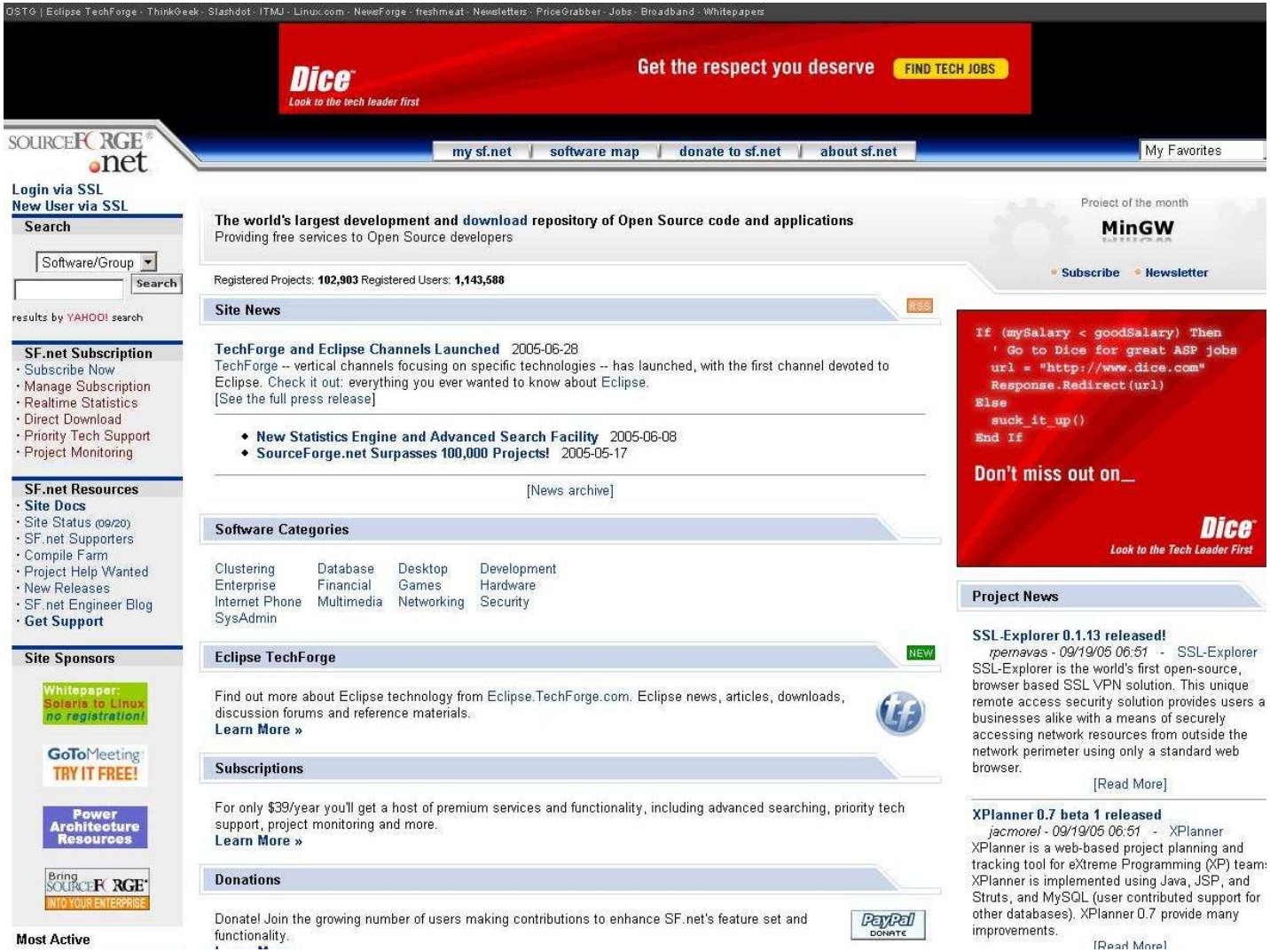
SOFTWARE PACKAGES

World Wind 1.3
The WorldWind virtual globe application and the independent components used to implement it.
Download Page

NASA Ames Research Center
Moffett Field, California 94035

Home | Sitemap | Feedback | FAQs | Other Centers
Life on Earth | Humans in Space | Exploring the Universe

SourceForge Site – http://www.sourceforge.net/



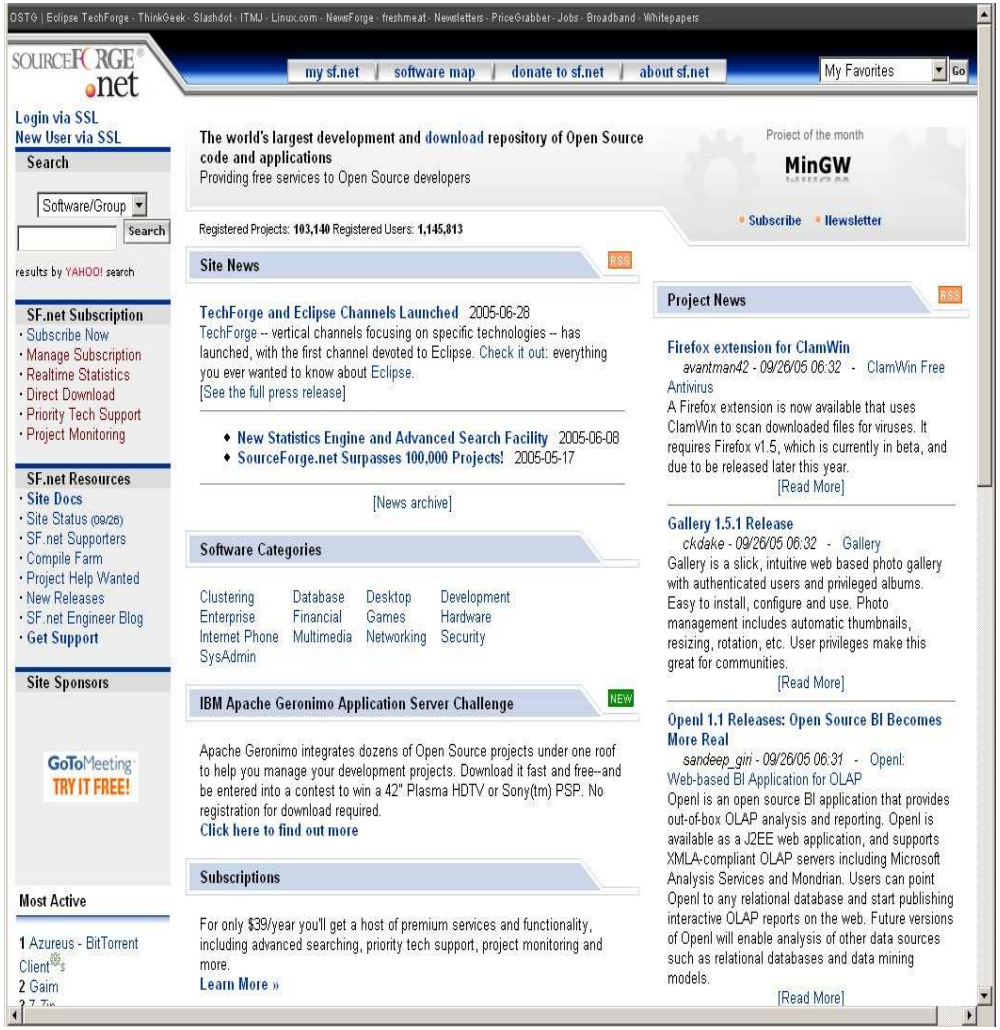
SourceForge Review Summary

- Domain is general software
- Type of assets provided are open source packages (approximately 103000)
- Repository
- Operational support is from eleven full-time staff members
- System technology is PHP web pages
- Site registration is available at levels (e.g., project admin)
- Assets can be added and updated
- System feedback available through tracker system; projects can offer other support
- Automatic notifications are available and based on subsets of project information
- Asset usage cannot be registered
- Asset reviews are not available
- System feedback is monitored
- Site does not have secure login/registration

SourceForge Review Table

Requirement/Feature	*Available at SourceForge?*
Domain	General software repository
Type of Assets	Open source packages
Register User	★★★
Contribute/Update Assets	★★★
Provide System Feedback	★☆☆
Automatic Notifications	★★☆
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★★☆
Secure Log In / Registration	YES
Catalog or Repository	Repository
Operation Support	Eleven full-time staff members
Technology	PHP

SourceForge Site – http://www.sourceforge.net/



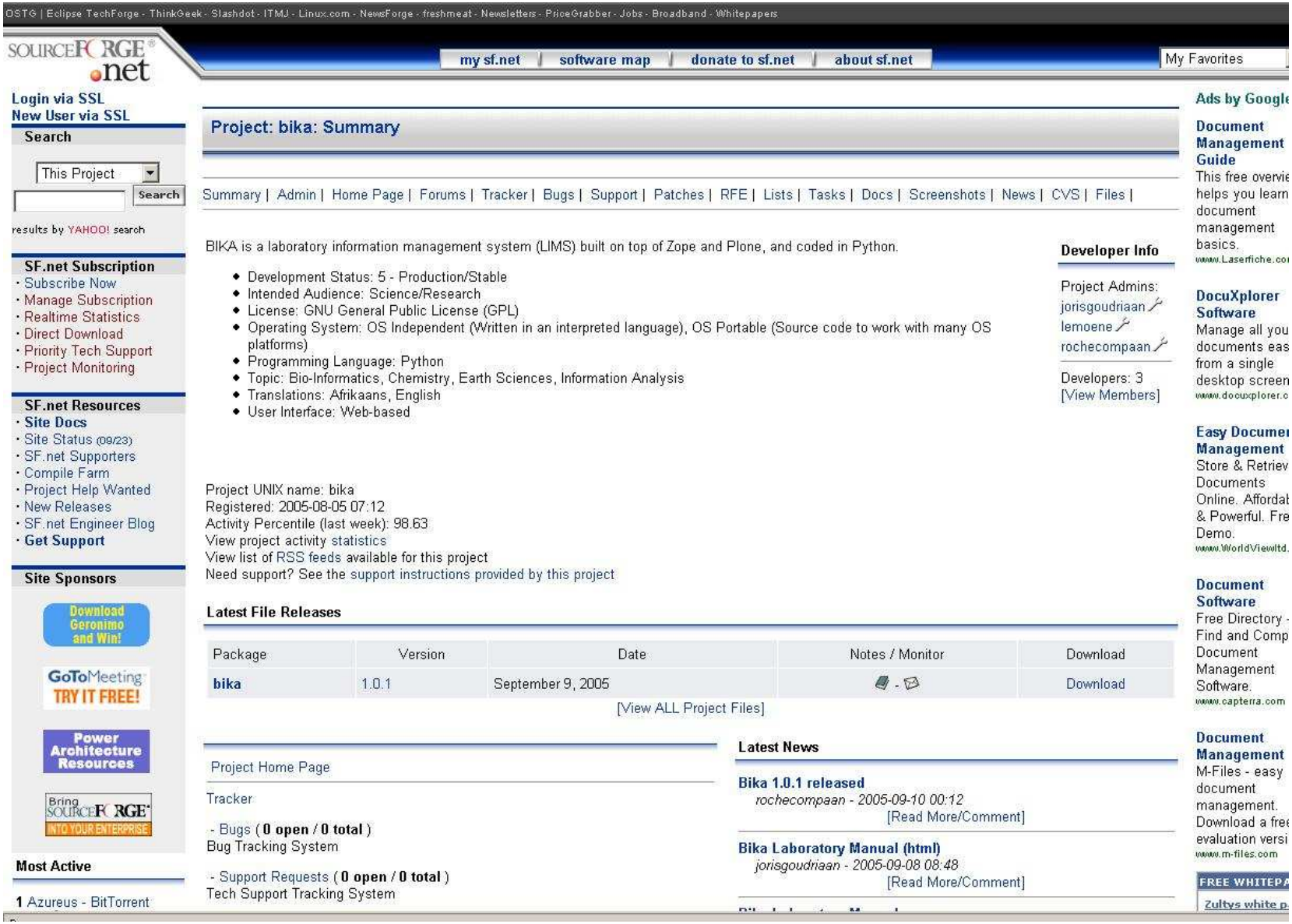
- Domain is general software
- Type of assets provided are open source packages (approximately 103000)
- Repository
- Operational support is from eleven full-time staff members
- System technology is PHP web pages

SourceForge Review

Requirement/Feature	*Available at SourceForge?*
Domain	General software repository
Type of Assets	Open source packages
Register User	★★★
Contribute/Update Assets	★★★
Provide System Feedback	★☆☆
Automatic Notifications	★★☆
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★★☆
Secure Log In / Registration	YES
Catalog or Repository	Repository
Operation Support	Eleven full-time staff members
Technology	PHP

- Site registration is available at levels (e.g., project admin)
- Assets can be added and updated
- System feedback available through tracker system; projects can offer other support
- Automatic notifications are available and based on subsets of project information
- Asset usage cannot be registered
- Asset reviews are not available
- System feedback is monitored
- Site does not have secure login/registration

SourceForge Project Page



NASA Systems Summary

Requirement / Feature	***Global Change Master Directory (GCMD)***	***GSFC Open Source Site***	***Ames Open Source Site***	***HDF-EOS Tools and Information Center***	***Computational Technologies Project***	***Earth Observing System Clearinghouse (ECHO)***	***Planetary Data Systems Software Download***
Domain	Earth science	Earth and space science	General science	Earth science, HDF/HDF-EOS	Earth and space science	Earth science	Planetary astronomy
Type of Assets	Data sets, data services	Open source packages	Open source packages	Applications	Applications and source code	Metadata	Tools, binaries and source
Register User	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	★☆☆	☆☆☆
Contribute/Update Assets	★★☆	★☆☆	★☆☆	☆☆☆	☆☆☆	★★☆	☆☆☆
System Feedback	★★☆	★★☆	★★☆	☆☆☆	★★☆	★★☆	★★☆
Automatic Notifications	★★★	★☆☆	★☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Discovering Assets	Hierarchy, Search	List	List	List, Filter	Hierarchy	Search	List
Register Asset Usage	☆☆☆	★★★	★★★	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Provide Asset Review	☆☆☆	★☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Monitoring Feedback	★☆☆	★★☆	★☆☆	☆☆☆	★☆☆	★☆☆	★☆☆
Secure Log In / Registration	N/A	NO	NO	NO	N/A	YES	N/A
Catalog or Repository	Catalog	Both	Both	Repository	Catalog	Catalog	Both
Operation Support	Large	Small	Small	Inactive	Small	Available	Small
Technology	RSYNC, Zope, CVS, Linux, Java, JavaServer Pages, XML, Apache, Oracle/PostgreSQL, Struts, Lucene, XSLT, Dreamweaver	PHP	JavaServer Pages	Cold Fusion	HTML	XML (WSDL), SOAP, UDDI	Cold Fusion

Non-NASA Systems Summary

Requirement / Feature	*Open Channel Foundation / COSMIC*	*SourceForge*	*Freshmeat*	*Scientific Applications on Linux*	*National Technology Transfer Center*	*National HPCC Software Exchange*	*Netlib*	*Savannah*	*Space Telescope Science Institute*	*Astronomical Software and Documentation Service*
Domain	General	General	General	Scientific	Federal technologies (mostly NASA)	HPPC	Mathematics	General	Astronomy	Astronomy
Type of Assets	Applications and source code	Open source applications	Open source applications	Tools and packages with source code	Applications	Tools and end packages	Source codes	Tools and packages	Packages, source	Packages, source
Register User	★★★	★★★	★★☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	★★☆	☆☆☆	☆☆☆
Contribute/Update Assets	★★☆	★★★	★★☆	☆☆☆	★☆☆	☆☆☆	★★☆	★★☆	☆☆☆	★★☆
System Feedback	★☆☆	★☆☆	★★☆	☆☆☆	★★☆	☆☆☆	★★☆	★★☆	★★☆	★★☆
Automatic Notifications	★★☆	★★☆	★★☆	☆☆☆	★★☆	☆☆☆	☆☆☆	★★☆	☆☆☆	☆☆☆
Discovering Assets	List, Hierarchy, Search	Hierarchy, Search	Hierarchy, Search	Hierarchy, Search (broken)	List, Hierarchy, Search	Hierarchy, Search	Hierarchy, Search	List, Search	List, Hierarchy	List, Hierarchy, Search
Register Asset Usage	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Provide Asset Review	☆☆☆	☆☆☆	★★★	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Monitoring Feedback	★★☆	★★☆	★★☆	☆☆☆	★☆☆	☆☆☆	★☆☆	★★☆	★☆☆	★☆☆
Secure Log In / Registration	YES	YES	NO	N/A	N/A	N/A	N/A	YES	N/A	N/A
Catalog or Repository	Repository	Repository	Repository	Catalog	Both?	Catalog	Repository	Repository	Repository	Catalog
Operation Support	Medium	Large	Medium	Inactive	Uncertain	Inactive	Large	Large	Small	Medium
Technology	PHP, MySQL	PHP	XML-RPC	HTML, Java	ASP	Repository In a Box	HTML	Perl, PHP, MySQL	HTML	HTML

Conclusions

- None of the existing sites fulfill the role of a software repository for the Earth science community.
- None of the systems that were evaluated sufficiently meet the requirements that are necessary to serve the community of Earth science software developers.
- Shortcomings of existing systems include the following:
  - Not meeting enough of the critical functional requirements
  - Not focusing on the Earth science domain
  - Not targeting software developers as the primary audience
  - Not providing the type of small-sized assets that are most desired by the community of Earth science software developers for reuse purposes

Conclusions

- A new catalog/repository system is needed to encourage and better enable software reuse within the community of Earth science software developers.
  - The GCMD is primarily a data provider and the system is not designed to be used as a software repository.
  - Similarly, ECHO is middleware acting as a data/service broker.
  - The Open Source Agreement sites have no real catalog or repository functionality and are restricted to NASA open source products.
  - Non-NASA sites are typically not domain-specific enough to meet the needs of a focused community.
- Some collaboration with existing systems may be possible, but existing systems alone cannot meet the needs of this community.

Recommendations

- NASA should establish an effective mechanism for dissemination of reusable assets within the Earth science community.
- NASA should evaluate the technology options for the provision of a reuse enablement system including:
  - commercial reuse catalogs/repositories
  - open source reuse catalogs/repositories
  - use of existing publicly available catalogs/repositories
  - custom build of a community-specific catalog
- Based on the conclusions of the technology evaluation, NASA should implement a reuse enablement system.
- NASA should develop guidelines and standards for the management and operation of a reuse enablement system.

Recommendations

- Impact for the Working Group
  - The reuse working group will evaluate the technology options for the provision of a reuse enablement system.
  - The reuse working group will develop guidelines and standards for the management and operation of a reuse enablement system.
  - The reuse working group will develop a proposal for the implementation of a reuse enablement system based on the conclusions of the technology evaluation.
- Desired Decision
  - HQ agreement to proceed with the evaluation of technology options and to provide funding for the evaluation.
  - HQ agreement in principle to the establishment of a reuse catalog/repository subject to the findings of the evaluation.